THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 267, No. 30, Issue of October 25, pp. 21936-21945, 1992
Printed in U.S.A.

Genomic Organization and Expression of the Human α_{1B}-Adrenergic Receptor*

(Received for publication, February 3, 1992)

Chodavarapu S. Ramarao, Julie M. Kincade Denker, Dianne M. Perez‡, Robert J. Galvin, R. Peter Riek, and Robert M. Graham§

From the Department of Heart and Hypertension Research, The Research Institute, Cleveland Clinic Foundation, Cleveland, Ohio 44195

α_1-Adrenergic receptors (ARs) are members of the guanine nucleotide-binding protein-coupled receptor superfamily. The genes for all ARs described thus far are intronless. We report here the cloning and the nucleotide sequence of the gene for the human α_{1B}-AR. It consists of two exons and a single large intron of at least 20 kilobases which interrupts the coding region at the end of the putative sixth transmembrane domain. The deduced amino acid sequence of the encoded receptor has a high degree of homology to the cloned hamster, rat, and dog α_{1B}-ARs. To characterize the encoded protein, a fusion gene constructed by splicing together exon 1 and exon 2 was expressed transiently in COS-1 cells. The transfected gene fusion product resulted in the production of an α_{1B}-AR with ligand binding characteristics indistinguishable from those of the expressed hamster α_{1B} cDNA. Evidence that the human α_{1B}-AR gene we have isolated is indeed transcribed is the finding of similar sized (2.8-kilobase) transcripts in human heart and other tissues by Northern blot analysis when either exon 1 or exon 2 is used as a probe. Moreover, using primers designed to span the exon 1/exon 2 boundary, a polymerase chain reaction product generated from single-stranded DNA prepared from human heart mRNA had the exact size and nucleotide sequence predicted for a transcript in which exon 1 is spliced to exon 2. The 5′-flanking region (924 base pairs (bp)) of exon 1 contains neither a TATA box nor a CAAT box but is high in GC content (70%) and contains several Sp1 binding sites (GC boxes), consistent with promoters described for housekeeping genes. The 5′-untranslated region also contains a putative cyclic AMP response element. Primer extension studies and RNase protection assays suggested that there are several potential transcription start sites in most tissues with a predominant site located 173 bp upstream from the translation start site. The 3′-flanking region contains a putative polyadenylation signal (ATTAAA) 492 bp downstream from the stop codon. The genomic organization of the human α_{1B}-AR with a single large intron interrupting its coding region differs from those of other ARs as well as muscarinic and 5-hydroxytryptamine receptors, which are intronless. The location of the intron in the human α_{1B}-AR gene is also unique among those members of the G-protein-coupled receptor family that do possess introns. Availability of this gene will now allow further studies on the transcriptional control of human α_{1B}-AR expression.

Adrenergic receptors belong to the superfamily of G-protein[1] coupled, seven transmembrane domain receptors. In response to external catecholamine stimuli, these receptors mediate a variety of cellular processes such as cardiac and arterial smooth muscle contraction and are thus involved in regulating cardiac function and in blood pressure homeostasis (1). ARs are broadly divided into α and β types based on their pharmacological specificities (2). Each class is further divided into several subtypes based on studies using both pharmacological and molecular cloning approaches. All of the members of the AR family described thus far are encoded by intronless genes. Thus genes for β_1- (3), β_2- (4), and β_3- (5) ARs and for various α_2-AR subtypes (6–8) lack introns. However, the genes for several other members of the G-protein-coupled receptor family contain introns. Genes encoding the rat substance P receptor (9), human tachykinin receptors (10, 11), dopamine D_2 (12) and D_4 (13) receptor subtypes, and opsins (14) contain introns at various locations in their coding or noncoding regions. The genomic organization of α_1-AR subtypes, however, has not been characterized. We report here the cloning of the gene encoding the human α_{1B}-AR. The intron/exon structure of the gene for this receptor is unique, and if shared by other α_1-AR subtypes, it may signify the existence of a distinct subgroup among ARs and other members of the G-protein-coupled receptor family.

EXPERIMENTAL PROCEDURES

Genomic Library Screening—Two different human genomic libraries were screened. One was a *Sal*I-digested human genomic DNA library constructed in EMBL3, and another was a partial *Sau*3A1-digested human genomic DNA library in λ-dash. Both libraries were screened with several degenerate oligonucleotide probes based on the published hamster sequence (15). Oligonucleotides were synthesized using a Cyclone DNA synthesizer, Milligen. The libraries were plated at a density of 50,000 plaques/plate. Approximately 1×10^6 plaques from each library were screened. Nylon membranes (Amersham Hybond N°) were lifted in duplicate from each plate, and the DNA was immobilized on the membranes by treating with denaturing solution and neutralizing solution followed by baking at 80 °C for 2 h in an oven. The membranes were probed with oligonucleotides end labeled

* This work was supported in part by National Institutes of Health Grants NS19583 and GM45985. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) M99589 and M99590.

‡ Recipient of National Research Service Award NS09056.

§ To whom correspondence and reprint requests should be addressed: Cleveland Clinic Foundation, FF3-20, 9500 Euclid Ave., Cleveland, OH 44195-5071. Tel.: 216-444-9262.

[1] The abbreviations used are: G-protein, guanine nucleotide-binding protein; AR, adrenergic receptor; bp, base pair(s); kb, kilobase(s); PCR, polymerase chain reaction; Tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine.

with [γ-^{32}P]ATP using T4 polynucleotide kinase. The hybridization conditions were 1×10^6 cpm/ml probe in 6 × SSC (1 × SSC = 0.3 M sodium chloride, 0.03 M sodium citrate buffer, pH 7.0), 0.1% SDS, and 0.7% sodium pyrophosphate, 1 × Denhardt's solution, and 0.1 mg/ml denatured herring sperm DNA. The filters were washed at 45 °C in 5 × SSC, containing 0.1% SDS and 0.7% sodium pyrophosphate. Several partial-length clones were isolated using these probes and purified to homogeneity by three rounds of screening. For identification of exon 2, the 3′ 220-bp and 480-bp *Pvu*II fragments of the hamster α_{1B}-AR cDNA were used as probes. The membranes were hybridized at 45 °C in the presence of 50% formamide, 2 × SSC, 0.1% SDS, 10% polyethylene glycol, 0.7% sodium pyrophosphate, 1 × Denhardt's solution, and 0.1 mg/ml denatured herring sperm DNA. The filters were washed at 55 °C in 2 × SSC, 0.1% SDS, 0.7% sodium pyrophosphate, and the positive clones were isolated by three rounds of purification. The positive clones were subjected to restriction map analysis, and the smallest fragments that hybridized to the probes by Southern blot analysis were subcloned into the pBluescript KS vector and sequenced in both directions using the dideoxy chain termination method (16). The complete nucleotide sequence was obtained both by subcloning and by using synthetic primers. Standard procedures were followed as described in Sambrook *et al.* (17).

cDNA Library Screening—A human heart λgt10 cDNA library was purchased from Clontech. The library was plated at a density of 50,000 plaques/plate, and duplicate membrane lifts were hybridized with probes labeled with [α-^{32}P]dCTP by random prime labeling using Klenow enzyme. Several probes were used at various screenings, either individually or in combination. The probes were derived from either exon 1 or exon 2 of the human genomic clones or were the 3′ *Pvu*II fragments of the hamster cDNA.

Construction of Exon 1/Exon 2 Gene Fusion Construct—A 2.5-kb *Eco*RI fragment bearing exon 1 and the 5′-flanking regions cloned in pGEM was used as a template for PCR amplification of exon 1 (see Fig. 8). The sense primer (39-mer) contained two *Eco*RI sites at the 5′ end followed by the Kozak consensus sequence for initiation of translation (18) and the first 21 nucleotides (underlined) from the start codon of the cloned human exon 1 (5′-gaattcgaattcgccaccatgaatcccgacctgcacacc-3′). The 34-mer antisense primer corresponded to the sequence at the 3′ end of exon 1 but without the final guanosine (5′-aagcggtagacgatgaagaagggtagccagcac-3′). The resulting PCR-amplified fragment was flush ended with murine leukemia virus reverse transcriptase, digested with *Eco*RI, and used further. A 2.7-kb *Bam*HI fragment, containing exon 2 and flanking regions, cloned in pBluescript KS, was used for the generation of exon 2. The plasmid was digested with *Pst*I, and the protruding 3′ ends were trimmed using the exonuclease activity of T4 DNA polymerase in the presence of (1 mM concentration each) dNTPs for 15 min at 15 °C. The linearized, blunt-ended plasmid DNA was digested with *Kpn*I. The resulting 787-bp fragment was isolated and used for constructing the fusion product. Modified exons 1 and 2 generated by the above procedures were ligated into the modified mammalian expression vector pMT2′ that has *Eco*RI and *Kpn*I sites, using a three-part ligation. The resulting plasmid was sequenced to verify the nucleotide sequences at the 5′- and 3′-cloning sites and at the exon 1/exon 2 splice site.

Transfection of COS-1 Cells—Plasmid DNA from the gene fusion construct purified by CsCl gradient centrifugation was used to transfect COS-1 cells (ATCC, Bethesda, MD) by the DEAE-dextran co-precipitation method as described (19). The pMT2′ plasmid without the insert was used for mock transfections. Membranes from COS cells transfected with the pMT2′ vector with or without the human α_{1B}-AR gene fusion construct, or with the α_{1B}-AR hamster cDNA, were prepared as described previously (20) and tested with various AR subtype selective and nonselective agonists/antagonists in competition binding assays, using the α_1-selective radioligand [^{3}H]prazosin. Details of the binding assays are as described (20).

Generation of End-specific cRNA Probes from Clones 1 and 2—Phage DNA obtained from clones 1 and 2 (see "Results" for the definition of clones 1 and 2) was digested to completion with *Rsa*I, dithiothreitol (0.3 M final concentration), 1 unit of RNase block II, and 50 μCi of [α-^{32}P]UTP were added to 1 μg of each digested template in a buffer containing 40 mM Tris, pH 8.0, 8 mM $MgCl_2$, 2 mM spermidine, and 50 mM NaCl. Ribonucleotides except UTP were added to a final concentration of 0.4 mM. The volume was made up to 25 μl with diethylpyrocarbonate-treated water. The reaction was initiated by addition of 10 units of T3 RNA polymerase (for clone 1) or T7 RNA polymerase (for clone 2). After incubating at 37 °C for 30 min, the DNA was digested with RNase-free DNase for 15 min at 37 °C. The reaction mixture was extracted with a mixture of phenol-chloroform, and the RNA was precipitated by 2.5 volumes of ethanol, in the presence of 0.3 M sodium acetate.

Primer Extension Analysis—To characterize the 5′ end of the mRNA, a ^{32}P end-labeled antisense oligonucleotide (corresponding to nucleotides 17–1) was annealed to 1 μg of poly(A$^+$) mRNA from four different human tissue sources (see Fig. 5). The mixture was heat denatured at 95 °C for 3 min, immediately transferred to 65 °C, and after 10 min, quenched on ice. The primer was extended with murine leukemia virus reverse transcriptase in the presence of deoxyribonucleotide triphosphates. The extension products were separated on a sequencing gel. The sizes of the resulting labeled, primer-extended products were inferred from their co-migration with a sequencing ladder, which was obtained using the same primer with the exon 1-containing clone.

RNase Protection Assay—An oligonucleotide duplex corresponding to nucleotides −213 to −148 was synthesized with *Eco*RI and *Bam*HI restriction sites at the ends and subcloned into pBluescript KS. Linearized plasmid DNA was used to generate radiolabeled cRNA using either T7 or T3 RNA polymerase, in the presence of ^{35}S-rATP. Total RNA from neuroblastoma and leiomyoma cell lines (100 μg) was hybridized with the cRNAs at 74 °C for 16 h in the presence of 4 M guanidinium isothiocyanate, 25 mM sodium citrate, and 0.5% sarcosyl in a volume of 50 μl. Unhybridized probe was digested by the addition of RNase A (20 μg/ml) and RNase T_1 (20 units/ml) in a volume of 500 μl containing 0.1 M NaCl. After a 30-min digestion at 37 °C, the mixture was treated with 10 μl of 20 mg/ml proteinase K, extracted with phenol and chloroform, precipitated with isopropyl alcohol, resuspended in formamide, and electrophoresed on a 6% sequencing gel.

Nested PCR Analysis of the 5′-Noncoding Region—PCR analysis of the 5′-noncoding region was performed using sets of nested primers. The primers used corresponded to nucleotides −1 to −21 (primer 1), −169 to −147 (primer 2), −192 to −173 (primer 3), −255 to −227 (primer 4), −557 to −603 (primer 5), −825 to −797 (primer 6), and −856 to −840 (primer 7). The templates used for these studies were either single-stranded DNA derived from random primed human heart poly(A$^+$) mRNA, or, as a control, a 2.5-kb *Eco*RI fragment of the gene cloned into pGEM that contains exon 1 and −1 kb of the 5′-flanking region. PCR was performed using the following conditions: 94 °C, 1 min; 60 °C, 1 min; and 72 °C, 1 min, for 40 cycles, followed by a 20-min extension period at 72 °C. The buffer used for PCR was 30 mM Tricine (pH 8.3), 1 mM $MgCl_2$, 5 mM β-mercaptoethanol, 0.01% gelatin, and 0.1% Thesit (21). PCR products were analyzed by electrophoresis using 2.5% Nusieve, 0.5% Seakem-agarose gels.

RESULTS

Cloning of the Human α_{1B}-AR Gene—To obtain the gene encoding the human α_{1B}-AR, two different human genomic libraries were screened with degenerate oligonucleotides based on the hamster α_{1B}-AR sequence (15). Using this approach, several clones were isolated from both libraries and were subjected to restriction analysis. A 2.5-kb *Eco*RI fragment from one of the clones was subcloned into pBluescript KS and sequenced in both directions. The sequence is highly homologous to the hamster α_{1B}-AR sequence but is interrupted in the coding region at the end of the putative sixth transmembrane region by an intron. To verify the existence of the intron, several positive clones from both libraries were sequenced partially at the exon/intron boundary using synthetic oligonucleotides. Since the location of the intron and the nucleotide sequence of the intron is identical among different clones isolated from two different libraries, it was concluded that the intron is not a cloning artifact. In the longest clone that was isolated, the intron was 13 kb (clone 1); however, this clone did not contain the rest of the coding region.

To obtain the missing 3′ sequence, the *Sau*3A1 genomic library was rescreened with degenerate oligonucleotide probes corresponding to the seventh transmembrane region of the hamster cDNA sequence. All of the positives obtained by this screening still had only exon 1. The reason for these clones

being recognized by the degenerate oligonucleotide probes was the high degree of similarity between the probes and short stretches of nucleotide sequences within the intron. Indeed, when the genomic DNA fragments from the clones that corresponded to the oligonucleotide probes were identified by Southern blot analysis, subcloned, and sequenced, the hybridizing sequences were still within the intron. To overcome this problem, the *Sau*3A1 library was screened again at moderate stringency (2 × SSC, 55 °C) with the 3′ 220- and 460-bp fragments obtained by digesting the hamster α_{1B}-AR cDNA with *Pvu*II. The two *Pvu*II DNA fragments extend from nucleotide positions 1020 to 1226 and from 1226 to 1686 of the hamster cDNA clone, respectively (15). The fragments represent the coding region 3′ to the sixth transmembrane region and thus should be homologous to the missing 3′ sequence of the human clone. Several clones were identified by this approach, and all of the clones were identical by restriction mapping and Southern blot analysis. A 2.7-kb *Bam*HI fragment thus identified was subcloned into pBluescript KS and sequenced in both directions. This fragment contained exon 2 as well as 1 kb of the intron at the 5′ end and 1 kb of sequence 3′ to the coding region. The longest clone identified in this manner has a 7-kb intron sequence upstream of exon 2 (clone 2).

Structure of the Human α_{1B}-AR Gene—Fig. 1 shows the genomic organization of the human α_{1B}-AR. The gene consists of two exons separated by a single long intron of at least 20 kb. The exon/intron boundary follows the consensus splice sequence AG/GT and is situated after the first base of the codon indicating a type I splice phase (22). The sequences at the donor and acceptor sites of the intron/exon boundaries match closely those of the consensus sequences (23). Exon 1 consists of the coding region that ends near the distal part of the putative sixth transmembrane domain. Exon 2 contains the rest of the coding region, as well as the 3′-noncoding region. There are two potential polyadenylation signals in the 3′-noncoding region. The sequences of the putative polyadenylation signals are identical to those found in the hamster. The putative signal TATAAA, starting 263 bases downstream from the stop codon, has one mismatch to the consensus AATAAA sequence. A second putative sequence, ATTAAA, starting 492 bases downstream from the stop codon, also has one mismatch to the consensus sequence. However, ATTAAA is the polyadenylation signal in the hamster (15), and chicken lysozyme genes (24), and in the mouse pancreatic α-amylase gene (25).

To determine whether the clones containing exon 1 and exon 2 overlap in their intronic regions, several experiments were performed. First, a 2.1-kb *Eco*RI-*Bam*HI fragment from the 3′ region of clone 1 was used as a probe to screen exon 2 containing clones. None of the exon 2-containing clones hybridized to this probe. Similarly, a 300-bp *Bam*HI fragment from the 5′ end of clone 2 failed to hybridize to any of the clones containing exon 1. In additional studies, ^{32}P-labeled cRNA probes derived from the 3′ end of clone 1 or the 5′ end of clone 2 were used to determine the presence of overlapping regions between the two sets of clones containing exon 2 and exon 1, respectively. Results from these experiments indicated that none of the exon 1-containing clones overlapped with the exon 2-containing clones. Finally, when the genomic library was rescreened with a part of the intron derived from clone 1, all of the positive reacting clones contained only exon 1. Rescreening of the genomic library with a fragment of the intron derived from clone 2 also identified only exon 2-containing clones.

Deduced Amino Acid Sequence of the Human α_{1B}-AR—Fig. 2 shows the nucleotide sequence and the deduced amino acid sequence of the human α_{1B}-AR. The sequence translates into a single 517-amino acid polypeptide with a molecular mass of 56,777 Da. The sequence is 2 amino acids longer than those of the hamster (15) and rat (26) α_{1B}-AR sequences. There are four sites for potential *N*-linked glycosylation at the amino terminus of the protein. A hydropathy plot indicates seven putative transmembrane domains (sequences *underlined* in Fig. 2), which are the hallmark of the G-protein-coupled receptor superfamily. As with several other members of the family, there is no signal peptide sequence in the molecule. The human α_{1B}-AR sequence is highly homologous to those of hamster, rat (26), and dog (27) sequences (Fig. 3). Comparison of the human α_{1B}-AR sequence with that of the hamster reveals 98% homology at the amino acid level and 89% identity at the nucleotide level. Comparison of the 3′-noncoding regions between the two sequences reveals a 76% homology. The differences in the amino acid sequences between the different species are mostly restricted to the carboxyl-terminal cytoplasmic tail, which for rhodopsin (28) and the β-AR (29), has been implicated in desensitization.

Transcription Initiation Site and Analysis of the 5′-Noncoding Region—The 5′-noncoding region of the human α_{1B}-AR gene contains nucleotide sequences consistent with donor (nucleotides −841 to −834) and acceptor (nucleotides −207 to −196) splice sites (Fig. 4*A*), suggesting the presence of a

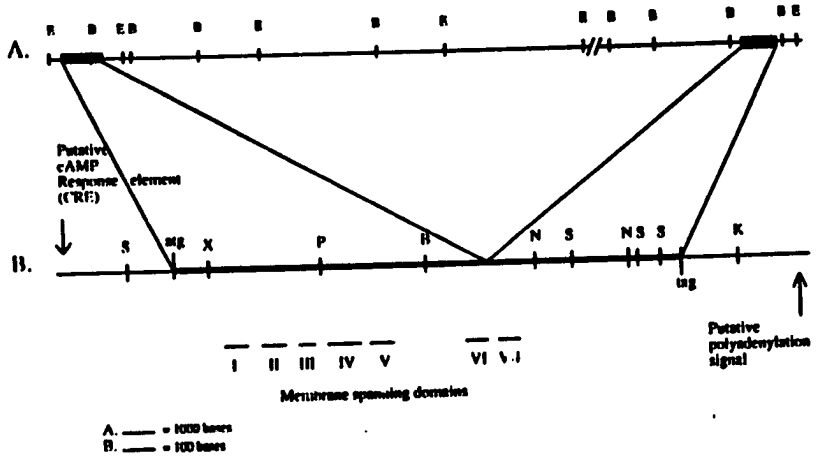

FIG. 1. Schematic of the genomic organization of the human α_{1B}-AR. *Panel A*, restriction map of the gene. The *line* representing the gene is interrupted at the location corresponding to the gap in the intron. Exons are indicated by *solid boxes*. Restriction sites are indicated by *vertical lines*. *Panel B*, detailed restriction map of the coding region. Putative cyclic AMP response element and polyadenylation sites are indicated. *B*, *Bam*HI; *E*, *Eco*RI; *K*, *Kpn*I; *N*, *Not*I; *P*, *Pst*I; *S*, *Sma*I; *X*, *Xho*I.

Fig. 2. Nucleotide and deduced amino acid sequence of the human α_{1B}-AR gene. Nucleotide numbering [illegible] with 1 at the translation initiation [illegible]. Nucleotide numbers are to the *right* [illegible]; Amino acid numbers are [illegible] nucleotide numbers. Coding [illegible] is shown in *upper case* letters, [illegible] noncoding sequences are [illegible] in lower case letters. Only [illegible] of intron sequence is shown. The size [illegible] is an approximation [illegible] for details). Sites for potential N-linked glycosylation are indicated [illegible]. Putative polyadenylation signals are *double underlined*. Predicted transmembrane domains are *underlined*. Amino acids are represented using the [illegible].

```
-63 ggagccttcgccgcagcccttccgagcccaatcatcccctggctatggagggcggactctaag ATG AAT CCC GAC CTG GAC ACC   21
                                                                Met Asn Pro Asp Leu Asp Thr    7
GGC CAC AAC ACA TCA GCA CCT GCC CAC TGG GGA GAG TTG AAA AAT GCC AAC TTC ACT GGC CCC AAC CAG ACC   93
Gly His Asn Thr Ser Ala Pro Ala His Trp Gly Glu Leu Lys Asn Ala Asn Phe Thr Gly Pro Asn Gln Thr   31
TCG AGC AAC TCC ACA CTG CCC CAG CTG GAC ATC ACC AGG GCC ATC TCT GTG GGC CTG GTG CTG GGC GCC TTC  165
Ser Ser Asn Ser Thr Leu Pro Gln Leu Asp Ile Thr Arg Ala Ile Ser Val Gly Leu Val Leu Gly Ala Phe   55
ATC CTC TTT GCC ATC GTG GGC AAC ATC CTA GTC ATC TTG TCT GTG GCC TGC AAC CGG CAC CTG CGG ACG CCC  237
Ile Leu Phe Ala Ile Val Gly Asn Ile Leu Val Ile Leu Ser Val Ala Cys Asn Arg His Leu Arg Thr Pro   79
ACC AAC TAC TTC ATT GTC AAC CTG GCC ATG GCC GAC CTG CTG TTG AGC TTC ACC GTC CTG CCC TTC TCA GCG  309
Thr Asn Tyr Phe Ile Val Asn Leu Ala Met Ala Asp Leu Leu Leu Ser Phe Thr Val Leu Pro Phe Ser Ala  103
GCC CTA GAG GTG CTC GGC TAC TGG GTG CTG GGG CGG ATC TTC TGT GAC ATC TGG GCA GCC GTG GAT GTC CTG  381
Ala Leu Glu Val Leu Gly Tyr Trp Val Leu Gly Arg Ile Phe Cys Asp Ile Trp Ala Ala Val Asp Val Leu  127
TGC TGC ACA GCG TCC ATT CTG AGC CTG TGC GCC ATC TCC ATC GAT CGC TAC ATC GGC GTG CGC TAC TCT CTG  453
Cys Cys Thr Ala Ser Ile Leu Ser Leu Cys Ala Ile Ser Ile Asp Arg Tyr Ile Gly Val Arg Tyr Ser Leu  151
CAG TAT CCC ACG CTG GTC ACC CGG AGG AAG GCC ATC TTG GCG CTG CTC AGT GTC TGG GTC TTG TCC ACC GTC  525
Gln Tyr Pro Thr Leu Val Thr Arg Arg Lys Ala Ile Leu Ala Leu Leu Ser Val Trp Val Leu Ser Thr Val  175
ATC TCC ATC GGG CCT CTC CTT GGG TGG AAG GAG CCG GCA CCC AAC GAT GAC AAG GAG TGC GGG GTC ACC GAA  597
Ile Ser Ile Gly Pro Leu Leu Gly Trp Lys Glu Pro Ala Pro Asn Asp Asp Lys Glu Cys Gly Val Thr Glu  199
GAA CCC TTC TAT GCC CTC TTC TCC TCT CTG GGC TCC TTC TAC ATC CCT CTG GCG GTC ATT CTA GTC ATG TAC  669
Glu Pro Phe Tyr Ala Leu Phe Ser Ser Leu Gly Ser Phe Tyr Ile Pro Leu Ala Val Ile Leu Val Met Tyr  223
TGC CGT GTC TAT ATA GTG GCC AAG AGA ACC ACC AAG AAC CTA GAG GCA GGA GTC ATG AAG GAG ATG TCC AAC  741
Cys Arg Val Tyr Ile Val Ala Lys Arg Thr Thr Lys Asn Leu Glu Ala Gly Val Met Lys Glu Met Ser Asn  247
TCC AAG GAG CTG ACC CTG AGG ATC CAT TCC AAG AAC TTT CAC GAG GAC ACC CTT AGC AGT ACC AAG GCC AAG  813
Ser Lys Glu Leu Thr Leu Arg Ile His Ser Lys Asn Phe His Glu Asp Thr Leu Ser Ser Thr Lys Ala Lys  271
GGC CAC AAC CCC AGG AGT TCC ATA GCT GTC AAA CTT TTT AAG TTC TCC AGG GAA AAG AAA GCA GCT AAG ACG  885
Gly His Asn Pro Arg Ser Ser Ile Ala Val Lys Leu Phe Lys Phe Ser Arg Glu Lys Lys Ala Ala Lys Thr  295
TTG GGC ATT GTG GTC GGT ATG TTC ATC TTG TGC TGG CTA CCC TTC TTC ATC GCT CTA CCG CTT Ggt aagtgg  949
Leu Gly Ile Val Val Gly Met Phe Ile Leu Cys Trp Leu Pro Phe Phe Ile Ala Leu Pro Leu G-          316
[illegible]gcagcaggg------------------> 20 kb <------------------cacccctccccactgcag
-GT TCC TTG TTC TCC ACC CTG AAG CCC CCC GAC GCC GTG TTC AAG GTG GTG TTC TGG CTG GGC TAC TTC AAC 1011
-ly Ser Leu Phe Ser Thr Leu Lys Pro Pro Asp Ala Val Phe Lys Val Val Phe Trp Leu Gly Tyr Phe Asn  340
AGC TGC CTC AAC CCC ATC ATC TAC CCA TGC TCC AGC AAG GAG TTC AAG CGC GCT TTC GTG CGC ATC CTC GGG 1084
Ser Cys Leu Asn Pro Ile Ile Tyr Pro Cys Ser Ser Lys Glu Phe Lys Arg Ala Phe Val Arg Ile Leu Gly  364
TGC CAG TGC CGC GGC CGC CGC CGA CGC CGC CGC CGC CGT CGC CTG GGC GGC TGC GCC TAC ACC TAC CGG CCG 1156
Cys Gln Cys Arg Gly Arg Arg Arg Arg Arg Arg Arg Arg Arg Leu Gly Gly Cys Ala Tyr Thr Tyr Arg Pro  388
TGG ACG CGC GGC GGC TCG CTG GAG CGC TCG CAG TCG CGC AAG GAC TCG CTG GAC GAC AGC GGC AGC TGC CTG 1228
Trp Thr Arg Gly Gly Ser Leu Glu Arg Ser Gln Ser Arg Lys Asp Ser Leu Asp Asp Ser Gly Ser Cys Leu  412
AGC GGC AGC CAG CGG ACC CTG CCC TCG GCC TCG CCG AGC CCG GGC TAC CTG GGC CGC GGC GCG CCA CCG CCA 1300
Ser Gly Ser Gln Arg Thr Leu Pro Ser Ala Ser Pro Ser Pro Gly Tyr Leu Gly Arg Gly Ala Pro Pro Pro  436
GTC GAG CTG TGC GCC TTC CCC GAG TGG AAG GCG CCC GGC GCC CTC CTG AGC CTG CCC GCG CCT GAG CCC CCC 1372
Val Glu Leu Cys Ala Phe Pro Glu Trp Lys Ala Pro Gly Ala Leu Leu Ser Leu Pro Ala Pro Glu Pro Pro  460
GGC CGC CGC GGC CGC CAC GAC TCG GGC CCG CTC TTC ACC TTC AAG CTC CTG ACC GAG CCC GAG AGC CCC GGG 1444
Gly Arg Arg Gly Arg His Asp Ser Gly Pro Leu Phe Thr Phe Lys Leu Leu Thr Glu Pro Glu Ser Pro Gly  484
ACC GAC GGC GGC GCT AGC AAC GGA GGC TGC GAG CCC CGC CAC GTG GCC AAC GGG CAG CCC GGC TTC AAA AGC 1516
Thr Asp Gly Gly Ala Ser Asn Gly Gly Cys Glu Pro Arg His Val Ala Asn Gly Gln Pro Gly Phe Lys Ser  508
AAC ATG CCC CTG GCC CCC GGG CAG TTT TAG ggcccccgtgcgcagctttctttccctggggaggaaaacatcgtggggtggaggg 1601
Asn Met Pro Leu Ala Pro Gly Gln Phe End                                                          517
[illegible]ccaaggggaacccgggggagggccggggagaggggcagctgcttt 1695
gagcgcggggcgcgggggagcggacggtcccacgacgggtagacgcgcgccaaggggaacc[illegible]gggatttcctc 1789
tctgccca[illegible]gggtgccaggtaccacgcgg[illegible]ccgggccgagcatgctgagagcctggggg[illegible]gtgcgtgtgttcc 1883
tctctctctctctgtgtatatataaacgcgtccctctctacttgtatttcaccgtgggtgcacgtgagtgtgactgtgcggtgtgcgtgtgttcc 1977
[illegible]gcccaggaggcaacc[illegible]gggcgttgtgtgtgtgtgactt[illegible] [illegible]
aag[illegible]ggatggcacttggtggggttggaacaaagtcgacattaaaggtcatttctcctgttcggcttt[illegible] [illegible]
[illegible]tgagcagagcct[illegible]tcctcttcctttg
```

[illegible] intron. Additionally, while this work was in progress, [illegible] reported that the human dopamine (D_{1A}) gene contains [illegible] the 5′ [illegible] coding region (23), although it lacks [illegible] intron in the translated region. To investigate if the human [illegible] AR gene contains an intron in the 5′-noncoding region, [illegible] was performed with various sets of primers, using either single-stranded DNA prepared from human heart mRNA or, [illegible] control, the 2.5-kb *Eco*RI fragment of the gene, which includes ~1 kb of the upstream region. The results of these [illegible] clearly indicate that the 5′-noncoding region of the [illegible] does not contain an intron. Thus, using single-stranded DNA prepared from mRNA, only PCR with primers 1 and 2, [illegible] but not 1 and 4, 1 and 6, 1 and 7, 5 and [illegible] 7 [illegible] yielded products of the expected [illegible] *Eco*RI gene fragment was used as [illegible] of these [illegible] of primers resulted in the [illegible] shows the PCR products [illegible] and 7 [illegible] in all instances, the identity [illegible] confirmed by restriction digestion. [illegible] single-stranded DNA yielded appropriate [illegible] and 8, but not with [illegible] start site [illegible] to primer 1 [illegible] that the 5′-noncoding region does not contain an intron. Additional evidence against the presence of an intron is the finding that PCR with single-stranded DNA and primers 1 and 7, and 1 and 6 did not yield products.

The transcription initiation site(s) of the human α_{1B}-AR message was determined by primer extension analysis using human poly(A^+) mRNA prepared from four different sources: brain, kidney, neuroblastoma (SK-N-M) cells (kindly provided by Dr. K. P. Minneman, Emory University, Atlanta), and leiomyoma (HS248.7) cells (ATCC). As shown in Fig. 5*A*, several potential starts were observed in all tissues except kidney, in which a single start site was observed. One of these sites located 173 bp upstream from the translation start site was clearly identified in all tissues tested (Fig. 5*A*).

To confirm the existence of multiple transcription initiation sites, RNase protection assays were performed. A synthetic oligonucleotide duplex corresponding to nucleotides −149 to −213 and containing *Eco*RI and *Bam*HI sites at the 5′ and 3′ ends, respectively, was subcloned into pBluescript KS. cRNA was then synthesized using T3 and T7 RNA polymerases in the presence of ^{35}S-rATP. Radiolabeled cRNA was allowed to hybridize to total RNA, and excess probe was digested with a combination of RNase A and T_1. The resultant protected cRNA probe was analyzed using a 6% urea polyacrylamide gel and autoradiography. Several RNA bands were visible as protected species (Fig. 5*B*), consistent with the existence of

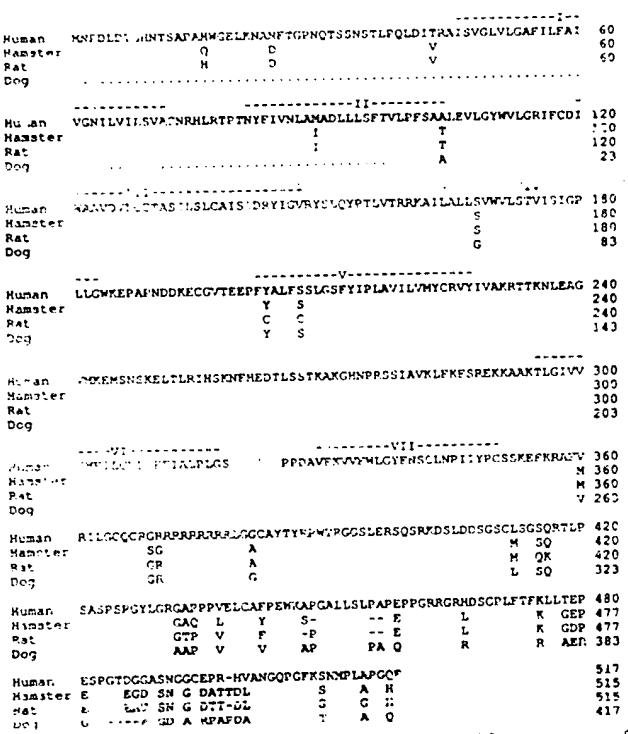

FIG. 3. **Alignment of the deduced amino acid sequences of the human, hamster (15), rat (26), and canine (27) α_{1B}-ARs.** Amino acids that are different from the human in at least one of the species are indicated. Gaps are introduced to optimize the alignment. Putative transmembrane domains are *overlined* and indicated by *Roman numerals*. The canine sequence is from a partial cDNA clone. The *dotted line* in the canine sequence indicates the missing amino-terminal sequence.

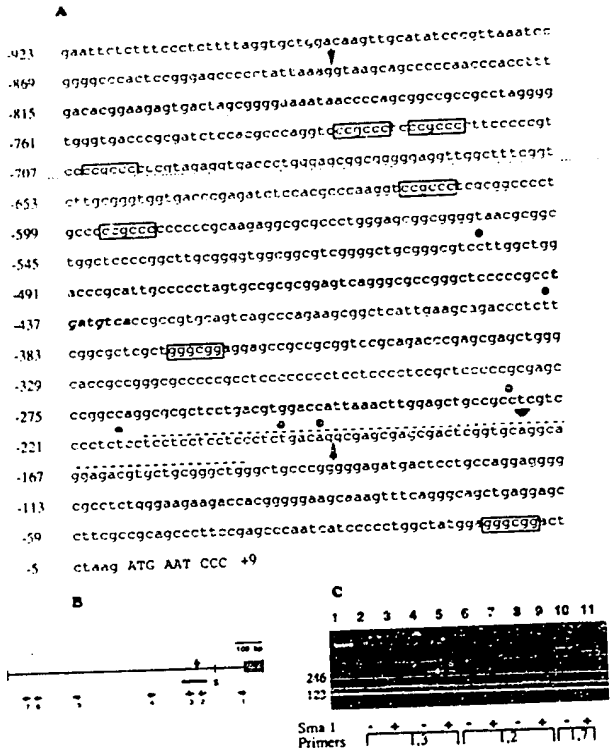

FIG. 4. **Analysis of the 5′-noncoding region of human α_{1B}-AR gene.** *Panel A*, nucleotide sequence of the 5′-noncoding region of the α_{1B}-AR gene. Numbers are relative to the adenosine of the initial ATG codon. Potential Sp1 binding sites are *boxed*. A potential consensus cyclic AMP responsive element is *italicized* in *bold*. Proposed major transcription initiation site is shown with a *solid triangle* above. Other transcription sites are indicated with *solid circles*. Boundaries of the potential intron are represented by *vertical arrows*. The nucleotide sequence of the synthetic duplex used in RNase protection experiments is *overlined* with a *dashed line*. *Panel B*, schematic of the analyzed region. The *horizontal line* represents the 5′-noncoding region. The *closed box* represents the coding region. The *vertical arrow* represents the major transcription initiation site determined by primer extension. All other transcription initiation sites determined by both primer extension and RNase protection assays are upstream to this position. The restriction site *SmaI*, used to characterize the PCR products, is shown as a *vertical line* with the letter *S* below it. The *heavy line* represents the region corresponding to the synthetic duplex used in RNase protection assays. *Horizontal arrows* with *numbers* represent the location and direction of primers used for PCR analysis. *Panel C*, analysis of PCR products. PCR was performed with different primer combinations as described under "Experimental Procedures." Plasmid DNA (containing the 2.5-kb *Eco*RI gene fragment that includes 1 kb of the 5′-flanking region) or single-stranded DNA derived from human heart poly(A$^+$) mRNA was used in the PCR reactions. Undigested (−) and SmaI-digested (+) PCR products were subjected to electrophoresis on 2.5% Nuseive, 0.5% Seakem-agarose gels. *Lane 1*, molecular weight standards (the positions of the 123- and 246-bp standards are indicated). *Lanes 2, 3, 6, 7, 10*, and *11*, PCR products obtained using plasmid DNA as template; *lanes 4, 5, 8*, and *9*, PCR products obtained using single-stranded DNA used as template. Primer combinations used in the PCR reactions are indicated.

multiple transcription initiation sites. The start sites determined by the RNase protection assay span nucleotides −177 to −198, when the migration of the protected products was compared with a sequencing ladder obtained using single-stranded DNA.

The 5′-untranslated region (924 bp) contains neither a TATA box nor a CAAT sequence. However, in this region, there are several potential Sp1 binding sites and a putative cyclic AMP response element (Fig. 4*A*). A putative cyclic AMP response element lies 264 bp upstream of the proposed major transcription initiation site and has the sequence TGA*TGTCA, which has a single base mismatch (indicated with an asterisk) to the consensus sequence (31).

A Single Transcript Is Recognized by Both Exon 1 and Exon 2—Northern blot analysis performed using the gene fusion construct identified an approximately 2.8-kb message in different tissues (Fig. 6). The size of this transcript is slightly larger than that of hamster message (2.6 kb) (15). Similar sized messages (2.8 kb) were also identified in human heart by Northern hybridization, when either exon 1 or exon 2 was used as probe (not shown).

The Human α_{1B}-AR Is Encoded by a Single Gene—Human genomic DNA was digested individually with eight different restriction enzymes, blotted, and probed with either exon 1 or exon 2. The number of fragments observed with [illegible] the α_{1B}-AR being encoded by a single-copy gene.

Expression of Human α_{1B}-AR in COS-1 Cells—Single-cell DNA could not be obtained despite repeated screening of a human heart cDNA library and since primary positive clones obtained from such screening did not generate PCR fragments of the expected sizes, we elected to splice together the two exons and express them in COS-1 cells to determine if a functional α_{1B}-AR is encoded by the gene fusion construct (Fig. 8). The inability to isolate a human dopamine D4 receptor cDNA by screening of several cDNA libraries and by using the PCR technique has been reported (13). Fusion of a partial cDNA with part of the gene to express a functional product has also been reported in the case of other AR receptors (32).

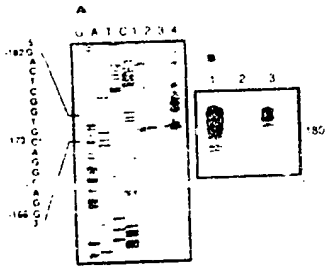

FIG. 5. **Determination of the transcription start site(s) of the human α_{1B}-AR gene.** *Panel A*, primer extension analysis. 1 μg of poly(A⁺) mRNA from human brain (*lane 1*), kidney (*lane 2*), leiomyoma cells (*lane 3*), or neuroblastoma cells (*lane 4*) was used in the analysis. A ^{32}P end-labeled antisense primer corresponding to nucleotides 17–1 of the human α_{1B}-AR gene was extended with murine leukemia virus reverse transcriptase. At the *left* is a sequencing reaction using the same primer and pGEM containing the 2.5-kb *Eco*RI fragment from clone 1. Several different potential transcription initiation sites are evident with all reactions. One site located 173 bp upstream from the translation start site, as indicated (*asterisk*), is common to all four tissues. The nucleotide sequence around this site is shown. *Panel B*, RNase protection assay. 100 μg of total RNA from neuroblastoma (*lane 1*) or leiomyoma (*lane 3*) was hybridized for 16 h at 74 °C to radiolabeled cRNA which was obtained by *in vitro* transcription of a synthetic oligonucleotide duplex corresponding to nucleotides −149 to −213. *Lane 2*, control hybridization without target RNA. The mixture was digested with RNases, and the resulting protected fragments were analyzed by electrophoresis on a 6% urea gel and autoradiography. The size of the protected fragments was determined from a simultaneously run DNA sequencing ladder, and the approximate migration of the fragments relative to markers 180 bp upstream from the initial ATG is shown.

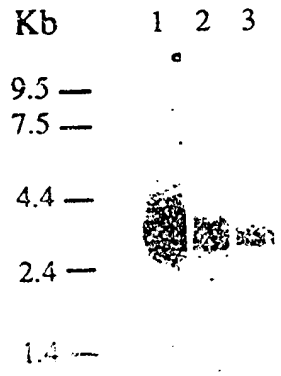

FIG. 6. **Northern blot analysis.** 2.5 μg of poly(A⁺) mRNA prepared from the human tissues indicated below was loaded per *lane*. The formaldehyde gel was washed twice with 10 × SSC and transferred to nylon membrane. RNA was fixed on the membranes by baking at 80 °C for 2 h. The blot was hybridized with a random prime-labeled exon 1 probe for 12 h at 43 °C, washed at 65 °C in presence of 0.2 × SSPE (1 × SSPE = 0.15 M NaCl, 0.01 M sodium phosphate, dibasic, 1 mM EDTA, pH 7.4), 0.2% SDS, and 0.7% sodium pyrophosphate and exposed to x-ray film. *Lane 1*, atrium; *lane 2*, kidney; *lane 3*, brain. *Numbers* at the *left* represent the RNA size markers.

When the α_{1B}-AR gene fusion construct was transfected into COS-1 cells, the cells expressed a functional receptor with ligand binding properties identical to those of the expressed human α_{1B}-AR (Table II). In addition, the rank order of potency of the adrenergic ligands that was observed with the expressed receptors indicates that the receptors are of the α_1 type (Table I) and not of the α_{1A}, α_{1C}, or α_{1D}-subtypes (for a review of the binding properties of these various α_1-AR subtypes see Ref. 20). Thus, for agonists the relative potency order is oxymetazoline > (−)norepinephrine > (+)norepinephrine > methoxamine, and for antagonists is prazosin > WB4101 > phentolamine > 5-methylurapidil > yohimbine > propranolol.

DISCUSSION

The gene that encodes the human α_{1B}-AR has been isolated and its structure determined. The coding region consists of two exons that are separated by a single large intron of at least 20 kb. In the absence of overlapping genomic clones, the exact size of the intron remains uncertain. Among the G-protein-coupled receptor superfamily, the rat substance P receptor gene has introns of comparable size. Two of the four introns in the rat substance P receptor are 15 and 23 kb (9). If the location of the introns is compared, the position of the intron in the human α_{1B}-AR gene is unique (Fig. 9). It appears that the intron/exon boundaries among the various G-protein-coupled receptor genes that contain introns are reasonably conserved among receptor subfamilies. All of the splice sites in the different opsins are conserved (Fig. 9). Several of the splice sites are also conserved among the genes for tachykinin receptors, as are those for the dopamine receptor genes that contain introns.

After repeated efforts, we were unable to isolate an α_{1B}-AR cDNA from a human heart cDNA library by either conventional screenings or by PCR amplification. Inability to isolate complete cDNAs has led others to develop cDNA-gene fusion constructs to express different G-protein-coupled receptors (13, 32). A 163-bp PCR fragment corresponding to the sequence from exon 1 was successfully isolated from the human heart cDNA library as well as from single-stranded cDNA prepared from human heart poly(A⁺) mRNA (not shown). This provides evidence that the human α_{1B}-AR gene is transcribed in the heart. However, mRNA corresponding to exon 2 was not amenable to PCR amplification, probably because of its high GC content and several imperfect direct repeats within the exon (Table II). In conventional screening using exon 1 from the human gene as a probe, several strongly positive clones could be identified in the primary screening of the cDNA library (20–30 positives/10^6 plaques) performed at very high stringency (0.1 × SSC, 70 °C wash). However, when these putative positive clones were plated for secondary screening, the signal intensity became very faint, and by the next round of purification, the hybridization signal was totally lost. Although it is difficult to ascribe the reason for this phenomenon, one possibility is the number of imperfect direct repeats found in exon 2 (Table II). Analysis of the human α_{1B}-AR coding sequence reveals that 366 bp out of the total coding length of 1,646 bp (or 23%) of the gene occurs as such repeats. Similarly, 33% of the human dopamine D_4 gene occurs as repeats, and neither traditional screening nor PCR was successful in isolating cDNA clones for this receptor (13). On the other hand, the rat α_{1D}-AR (20) has only 2.7% of the sequence as imperfect direct repeats. Palindromic and other repetitive sequences have been shown to be lethal for several cells and are often removed by deletion, forming cruciform structures, or by a DNA slippage mechanism during homologous recombination events (33). Repetitive DNA sequences, which can potentially form hairpin loops, are suggested to play a role in gene regulation (34) and in the termination of transcription initiated by RNA polymerase II (35). Repeated DNA sequences are also implicated in the retardation of growth of bacteriophages (36) and plasmids (37). This could explain, at least in part, our inability to isolate the cDNA clone by conventional screening or by PCR from the library

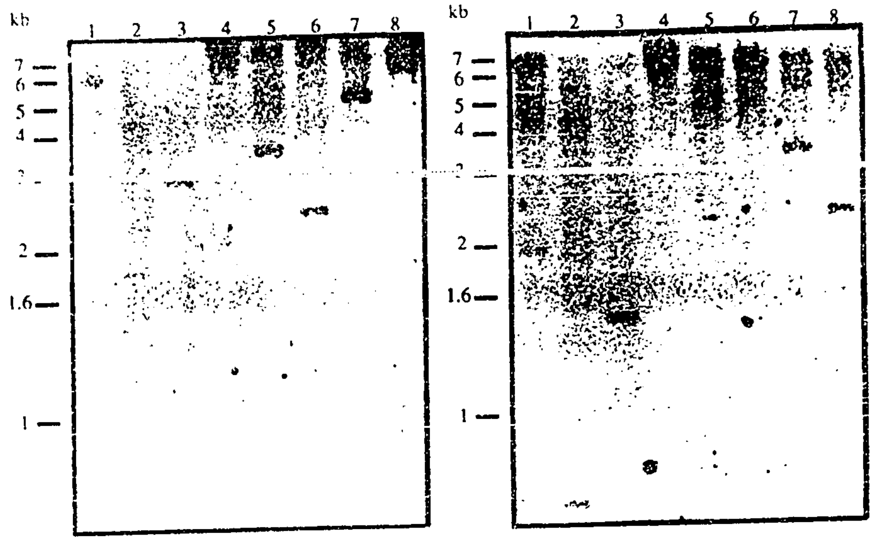

FIG. 7. **Southern blot analysis of the human α_{1B}-AR gene.** 10 μg of human genomic DNA was digested with the [illegible] ... probes were used: (i) a 7[illegible]-bp *Xho*I-*Bam*HI fragment derived from exon 1 (*left panel*), and (ii) a 1.7-kb *Pst*I fragment that included exon 2 and the 3′-noncoding region (*right panel*). *Lane numbers* indicate the restriction enzymes used to digest the DNA. *Lane 1*, *Sac*I; *lane 2*, *Pvu*II; *lane 3*, *Pst*I; *lane 4*, *Kpn*I; *lane 5*, *Hin*dIII; *lane 6*, *Eco*RI; *lane 7*, *Bgl*II; *lane 8*, *Bam*HI. *Numbers* on the *left* refer to the DNA size markers.

that apparently contained the clone of interest. Alternatively, sequences in the 5′- and/or 3′-noncoding regions may be responsible for our inability to isolate the cDNA clones since the plasmid containing the gene fusion construct, consisting of only the coding region, could be readily propagated in bacteria.

In the absence of a cDNA for the α_{1B}-AR, we have developed a number of different lines of evidence to indicate that exon 1 and exon 2 are indeed part of the same gene. First, there is the high degree of homology between the human sequence and the hamster, rat, and canine α_{1B}-AR sequences. In fact, even the 3′-noncoding regions of the human and hamster are conserved with 76% identity at the nucleotide level. It is believed, based on comparisons of members of a gene family, and of a given gene in different species, that although coding regions evolve slowly, the noncoding sequences (*i.e.* introns and very often 3′-noncoding sequences) evolve much more rapidly (23). Second, expression of the gene fusion construct in COS cells produced a functional α_{1B}-AR. The expressed protein has a pharmacological profile that is essentially indistinguishable from that of the hamster α_{1B}-AR clone (Table I). Third is the identification of similar sized transcripts in human heart and other tissues by both exon 1 and exon 2 probes. Fourth, the intron/exon boundaries are likely correct since the sequences in the region match closely with the consensus sequences around the donor and acceptor splice sites. The nucleotide sequence around the donor site T*G/GTAAGT (where the slash represents exon/intron boundary) has a single base mismatch (indicated with an asterisk) to the consensus sequence AG/GTAAGT (22, 37). Similarly, the sequence near the acceptor site CCCTCCCCA*CTGCAG/G (where the slash represents the intron/exon border) also has a single base mismatch to the consensus sequence $(C/T)_{11}$ NCAG/G (38). The extreme 3′ end of the intron is pyrimidine rich (82%) which is a general feature of the sequences near the end of an intron (23, 38, 39). Furthermore, the branch point signal sequence TACTAAC occurring 20–50 bases upstream from the 3′ splice site, which is highly conserved in yeast (38) and weakly conserved in mammals (39), is represented in the human α_{1B}-AR gene with a single base mismatch, TACTC*AC, and occurs 52 bases upstream from the 3′ splice site. Finally, the strongest evidence that the two exons form a single message comes from studies based on PCR amplification (Fig. 10). In these studies, single-stranded DNA prepared from human heart mRNA was used as a template. The primers used corresponded to the sequences on either side of the splice site. The sense primer (5′-gaattcgaattccacaaccccaggagttccatagctgtcaaactttttaag-3′) (*Eco*RI sites were added at the 5′ end to facilitate subcloning, and are underlined) corresponded to nucleotide sequence 817–855 (Fig. 2). The antisense primer extended from nucleotide position 1102 to 1063 (5′-aagcttaagcttgcactggcaccccgaggatgcgcacgaaagcgcgcttgaa-3′) (*Hin*dIII sites were added to the 5′ end and are underlined). The 284-bp product obtained by PCR was subcloned and sequenced to reveal an identical nucleotide sequence at the end and start of exons 1 and 2, respectively (Fig. 10).

The lack of a TATA box upstream from the origin of transcription is not unusual. Several genes are known in which TATA boxes are not present in the upstream transcription initiation site. Such TATA-less genes can be divided into two classes. One class comprises housekeeping genes containing mainly GC-rich promoters (40), with several transcription initiation sites spread over a large region and having several potential Sp1 binding sites (41, 42). The second class have neither TATA boxes nor GC boxes, and they are not constitutively active (41). The nucleotide sequence around the putative transcription initiation site in the human α_{1B}-AR gene is rich in GC content, and there are several consensus Sp1 binding sites. Primer extension analysis suggested several transcription initiation sites with a predominant site located 173 bp upstream from the translation initiation site (Fig. 5*A*). Multiple transcription initiation sites were also identified by RNase protection assay (Fig. 5*B*). In several respects the 5′-noncoding region of human α_{1B}-AR gene resembles the recently characterized dopamine D_{1A} gene (30). Similarities between the two genes include a high GC content of the 5′-noncoding region, the lack of a TATA box and CAAT box, the presence of several Sp1 binding sites and multiple transcription initiation sites, and the presence of a single intron. However, in the D_{1A} receptor the intron is in the 5′-noncoding region, whereas in the α_{1B}-AR gene the intron is in the coding region and is very long. Although there are consensus sequences for donor and acceptor splice sites, which may indicate the presence of an additional intron in the 5′ region of the α_{1B}-AR gene, the RNase protection, primer extension, and PCR studies using nested primers showed conclusively that there is no additional intron in this region. Analysis of the 5′ end of the β_2-AR gene (4) revealed putative TATA boxes that are poorly conserved. The closest approximation to a TATA box in the human β_2-AR gene has been reported

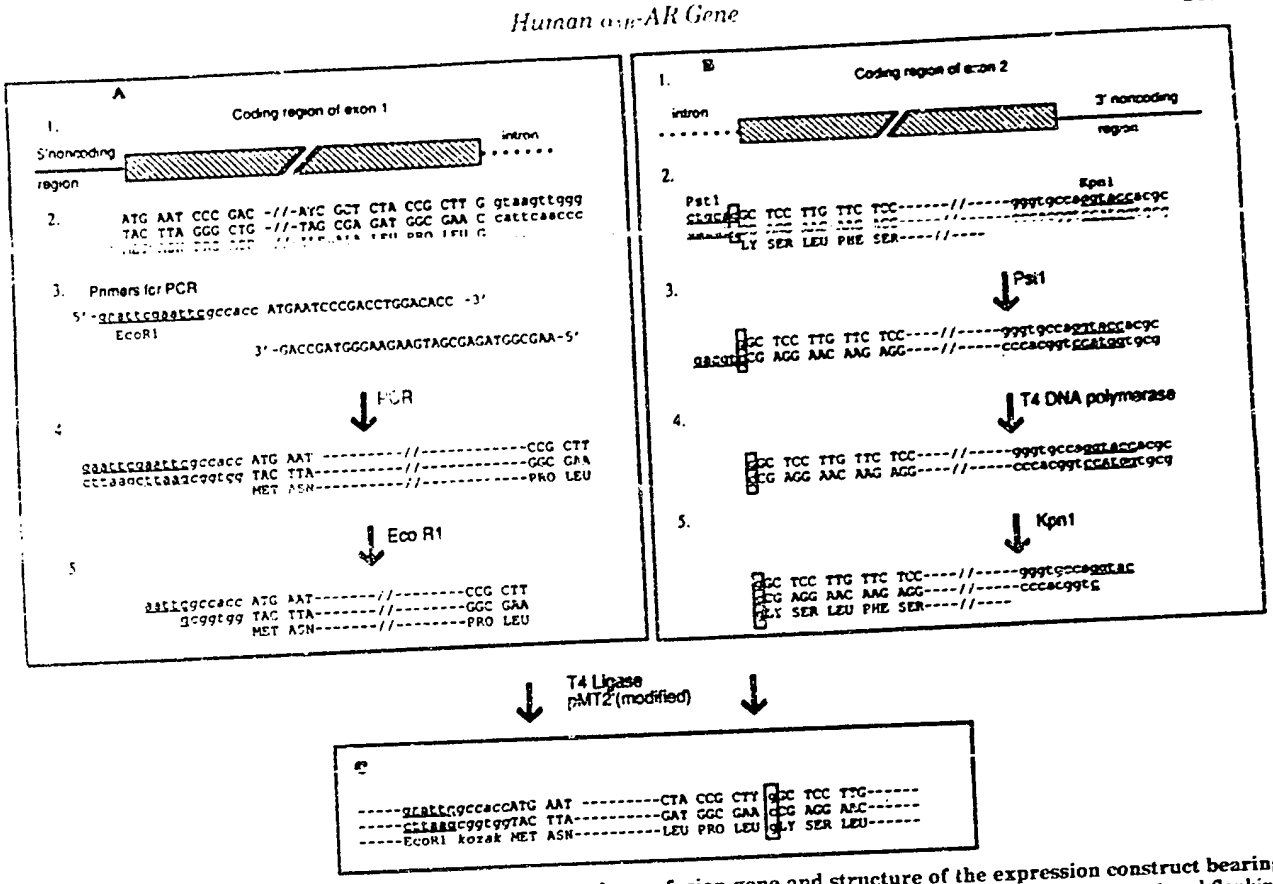

FIG. 8. **Strategy for splicing exon 1 and exon 2 to produce a fusion gene and structure of the expression construct bearing the human α_{1B}-AR gene.** *Panel A*, strategy used to obtain exon 1. *1*, a 2.5-kb *Eco*RI fragment from clone 1 containing exon 1 and flanking regions was cloned into pBluescript KS and used as a template for PCR. The *hatched box* represents the coding region of exon 1. The intron is represented by the *dotted line*, and the 5′-noncoding region is shown as a *solid line*. *2*, below the *box* is a partial sequence of the coding region at the translation initiation site and splice site. The coding region is represented in *upper case letters*, and the intron sequences are represented in *lower case letters*. The corresponding deduced amino acid sequence is shown. *3*, primers used for PCR. The 5′ end of the sense primer has two *Eco*RI sites (*underlined*) and a Kozak consensus sequence (*italicized*) for translation (18). The antisense primer lacks the ultimate guanine nucleotide. PCR was performed using the following profile: 94 °C, 1 min; 60 °C, 1 min; 72 °C, 1 min for 40 cycles followed by 72 °C for 20 min. *4*, structure of the PCR-amplified product. *5*, the PCR product was digested with *Eco*RI. *Panel B*, strategy used to obtain exon 2. *1*, schematic of the 2.7-kb *Bam*HI fragment containing exon 2 and flanking regions cloned into pBluescript. Notations are similar to those under *A1*. The G-C base pair derived from the intron that will become part of the coding region is *boxed*. *2*, partial sequence of the exon and 3′-noncoding region showing the unique restriction sites *Pst*I and *Kpn*I used in the construction. *3*, the DNA was digested with *Pst*I. *4*, the resulting 3′ overhang was converted to a blunt end using the exonuclease activity of T4 DNA polymerase. *5*, the linearized plasmid DNA was digested with *Kpn*I, and the resultant fragment was ligated with exon 1 obtained in *A5* and with the modified pMT2′ vector to yield the fusion product (*panel C*). Note that the the G-C base pair at the splice site is now derived from the intron.

to be either ATAA or CTAA (4, 5). The consensus sequence CTCANTCT near the start of transcription can replace the TATA box (41, 43). However, we do not find any sequence to match this consensus motif at an appropriate position upstream of the transcription site of the α_{1B}-AR.

The intronless nature of several G-protein-coupled receptors has greatly aided in the isolation of various AR subtypes. G-protein-coupled receptors have several functional/structural domains like transmembrane segments, a ligand binding site, G-protein interacting domains, and domains that are involved in down-regulation/desensitization. In several cases, intron/exon separation occurs at or near the putative transmembrane boundaries (44). Thus, the intron/exon organization might be useful in separating the protein into discrete functional domains (44), and splicing could underlie receptor subtype diversity, as reported for the dopamine D_2 receptor (45). Considering the diversity of G-protein-coupled receptors on the one hand and the diversity of various G-protein subunits themselves on the other (45), it is surprising that most of the AR genes are intronless. By possessing a single large intron at the end of the putative sixth transmembrane domain, the α_{1B}-AR gene differs from those of the other AR subtypes. A comparison of the intron/exon organization of the family members that do have introns might be helpful in tracing the evolutionary relationship between the various members. As a first step, it could be concluded that based on the genomic organization, the α_{1B}-AR (and possibly other α_1-AR subtypes) are encoded by genes that are more distantly related to the rest of the AR family. It is interesting to speculate that α_1-AR receptors differ from other ARs in that they are encoded by genes that have introns. Indeed, preliminary studies indicate that the human α_{1C}-AR gene has an intron at the same location as that observed here for the α_{1B}-AR gene.[2] Thus, one can predict, from an evolutionary standpoint, that α_1-AR subtypes form a distinct subfamily. And indeed, based on sequence homologies, the α_{1B}-AR has been separated from the α_2- and β-ARs using a hypothetical cationic amine receptor dendrogram (46).

In summary, we have cloned and characterized the gene

[2] J. Tseng-Crank, personal communication

TABLE I

Pharmacological characterization of the expressed human α_{1B}-AR

COS-1 cells were transfected with pMT2′ vector containing the human α_{1B}-AR gene fusion construct (Fig. 6) or the hamster α_{1B}-AR cDNA. Mock tranfections were performed with the vector without the insert. Membranes prepared from the transfected cells were incubated with the α_1-AR antagonist [^{3}H]prazosin in the presence or absence of increasing concentrations of various antagonists or agonists. Results shown are the K_i values determined from the competition studies as described (20). Values shown are the means of at least two experiments with each ligand performed in duplicate.

Ligands	Human α_{1B} gene fusion construct	Hamster α_{1B}-AR
Agonists (μM)		
(−)Norepinephrine	1.[illegible]	1.9
(+)Norepinephrine	106	131
Oxymetazoline	0.11	0.12
Methoxamine	495	619
Antagonists (nM)		
Prazosin	0.035	0.03[illegible]
WB4101	[illegible].00	6.97
Phentolamine	[illegible]1	39
5-Methylurapidil	64	70
Yohimbine	1,562	1,003
(−)Propranolol	15,690	9,190

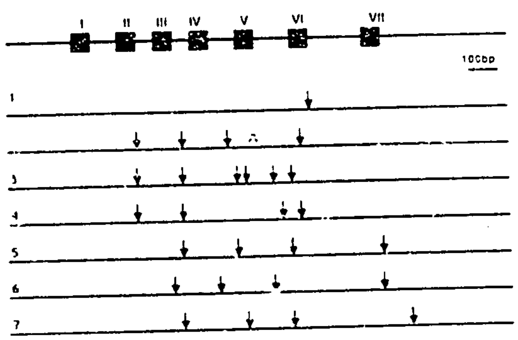

FIG. 9. **Comparison of the intron/exon boundaries of some of the G-protein-coupled receptors that are encoded by genes containing introns.** The human α_{1B}-AR gene is compared with other G-protein-coupled receptor genes containing introns, including the rat substance P receptor (9), human opsins (14), human dopamine D_4 (13) and D_2 (12) receptors, and human tachykinin receptors (10, 11). Transmembrane domains are shown as *closed boxes*. Intron/exon boundaries are indicated with *arrows*. The genes represented are: *1*, α_{1B}-AR; *2*, rat dopamine D_2 receptor; *3*, rat dopamine D_3 receptor; *4*, human dopamine D_4 receptor; *5*, rat substance P receptor; *6*, human opsin; and *7*, human neurokinins 1 and 2. The *large arrow* in the dopamine receptor gene indicates the presence of an alternate splice site.

TABLE II

Direct imperfect repeats of nucleotide sequences found in the coding region of the human α_{1B}-AR gene

The human α_{1B}-AR coding sequence was analyzed using the Compare program in Microgenie to identify direct repeats. The parameters used to define direct repeats were a minimum length of 25 nucleotides at a minimum of a 70% match. No inverted repeats were found at the same stringency.

Nucleotides	Repeated	No. of bases in repeat	% match in repeat
[illegible]	617–644	31	7[illegible]
[illegible]–712	1393–1421	32	78
71[illegible]–74[illegible]	1008–1037	30	73
11[illegible]–1207	1371–1397	27	78
1[illegible]–146[illegible]	1473–1499	28	75
[illegible]	1516–1560	16	75

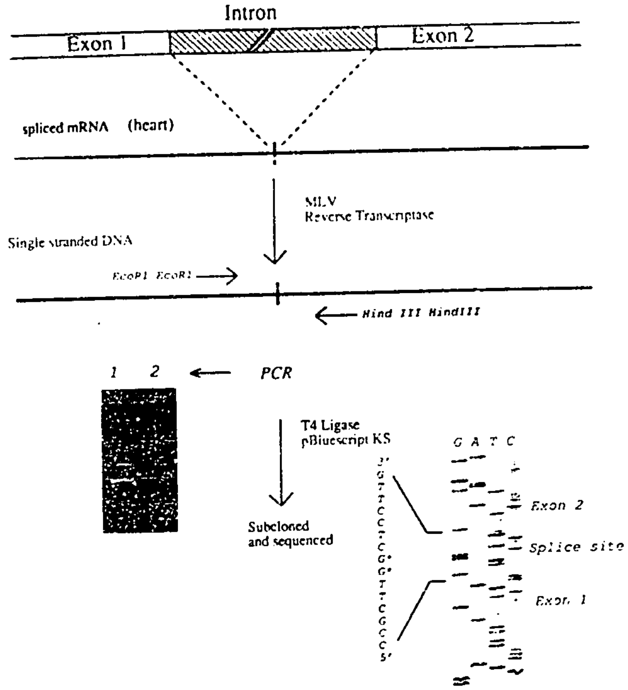

FIG. 10. **Exons 1 and 2 of the human α_{1B}-AR gene encode a single mRNA.** A schematic of the gene is shown at the *top*. The *discontinuous hatched box* represents the intron. Processing the mRNA (*horizontal line*) involves splicing together the two exons after removal of the intron. The splice site is shown as a *vertical line*. Poly(A$^+$) mRNA from human heart was used to make single-stranded DNA using murine leukemia virus reverse transcriptase and an oligo(dT) primer. The resulting single-stranded DNA was then used for PCR with two synthetic oligonucleotide primers corresponding to the exon sequences on either side of the splice site (see "Discussion" for details). The resulting PCR product (*bottom left*: *lane 1*, DNA size marker; *lane 2*, 294-bp PCR product) was subcloned into pBluescript and sequenced to identify the splice site (indicated with *asterisks* at *bottom right*). To rule out the possibility of contamination during PCR with the gene fusion construct contained in pMT$_2$′ (Fig. 8), different combinations of pMT$_2$′-specific primers and α_{1B}-AR specific primers were used in parallel experiments with either single-stranded DNA obtained from mRNA, or with the gene fusion construct. Only the expected sized products were obtained with each set of nested primers, thus excluding contamination.

that encodes the human α_{1B}-AR. The gene is unique among the members of the AR family by virtue of a single long intron that separates the coding region at the end of the sixth transmembrane region. A gene fusion product constructed with exons 1 and 2, when expressed in COS-1 cells, produces a functional α_{1B}-AR. The deduced amino acid sequence of the human α_{1B}-AR is highly homologous to those of hamster, rat, and canine α_{1B}-ARs. The gene apparently lacks a TATA box but possesses several potential Sp1 binding sites around the proposed start site of transcription. Characterization of this gene will now permit studies on the regulation of α_{1B}-AR expression.

Acknowledgments—We thank C. A. Boehm for mRNA preparation and for tissue culture and Dr. S. Karnik for valuable discussions. We also thank J. Holl for typing the manuscript. The hamster α_{1B}-AR cDNA (15) was kindly provided by Dr. C. E. Seidman, Harvard Medical School. We thank Dr. J. Tseng-Crank for sharing unpublished data on the human α_{1C}-AR gene structure.

REFERENCES

1. Raymond, J. R., Hnatowich, M., Lefkowitz, R. J., and Caron, M. G. (1990) *Hypertension* **15**, 119–131
2. Ahlquist, R. P. (1948) *Am. J. Physiol.* **153**, 586–600

3. Dixon, R. A. F., Kobilka, B. K., Strader, D. J., Benovic, L. J., Dohlman, H. G., Frielle, T., Balanowski, M. A., Bennet, C. D., Rands, E., Diel, R. E., Mumford, R., Slater, E. E., Sigal, I. S., Caron, M. G., and Lefkowitz, R. J. (1986) *Nature* **321**, 75–79
4. Kobilka, B. K., Frielle, T., Dohlman, H. G., Balanowski, M. A., Dixon, R. A. F., Keller, P., Caron, M. G., and Lefkowitz, R. J. (1987) *J. Biol. Chem.* **262**, 7321–7327
5. Emorine, L. J., Marullo, M. M., Rieno-Sutren, B., Patery, G., Kluckke, K. T., and Strosburg, A. D. (1989) *Science* **245**, 1118–1121
6. Weinshank, R. L., Zgombrik, J. M., Macchi, M., Adham, N., Litchblau, H., Branchek, T. A., and Hartig, P. R. (1990) *Mol. Pharmacol.* **38**, 681–688
7. Kobilka, B. K., Matsui, H., Kobilka, T. S., Yang Feng, T. L., Franck, U., Caron, M. G., Lefkowitz, R. J., and Regan, J. W. (1987) *Science* **238**, 650–656
8. Lomasney, J. W., Lorenz, W., Allen, L. F., King, K., Regan, J. W., Feng, T. L., Caron, M. G., and Lefkowitz, R. J. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 5094–5098
9. Hershey, A., Dykema, P. E., and Krause, J. E. (1991) *J. Biol. Chem.* **266**, 4366–4374
10. Gerard, N. P., Eddy, R. L., Jr., Shows, T. B., and Gerard, C. (1990) *J. Biol. Chem.* **265**, 20455–20462
11. Gerard, N. P., Garraway, L. A., Eddy, R. L., Jr., Shows, T. B., Iijima, H., Paquet, J. L., and Gerard, C. (1991) *Biochemistry* **30**, 10640–10646
12. Toso, R. D., Sommer, B., Ewert, M., Herb, A., Pritchett, D. B., Bach, A., Shivers, B. D., and Seeburg, P. H. (1989) *EMBO J.* **8**, 4025–4034
13. Van Tol, H. H. M., Bunzow, J. R., Guan, H. C., Sunahara, R. K., Seeman, P., Niznik, H. B., and Civelli, O. (1991) *Nature* **350**, 610–614
14. Nathans, J., Thomas, D., and Hogness, D. S. (1986) *Science* **232**, 193–202
15. Cotecchia, S., Schwinn, D. A., Randall, R. R., Lefkowitz, R. J., Caron, M. G., and Kobilka, B. K. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 7159–7163
16. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
17. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Vol. 1–3, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. Kozak, M. (1987) *Nucleic Acids Res.* **15**, 8125–8148
19. Kauffman, R. (1990) *Methods Enzymol.* **185**, 487–511
20. Perez, D. M., Piascik, M. T., and Graham, R. M. (1991) *Mol. Pharmacol.* **40**, 876–883
21. Ponce, M. R., and Micol, J. L. (1992) *Nucleic Acids Res.* **20**, 623
22. Rogers, J. (1985) *Nature* **315**, 458–459
23. Breathnach, R., and Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383
24. Jung, A., Sippel, A. E., Grez, M., and Schutz, G. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 5759–5763
25. Hagenbuchle, O., Bovey, R., and Young, R. A. (1980) *Cell* **21**, 179–187
26. Voight, M. M., Kispert, J., and Chen, H. (1990) *Nucleic Acids Res.* **18**, 1053
27. Libert, F., Parmentier, M., Lefort, A., Dinsart, C., Van Sande, J., Meanhaut, G., Simons, M.-J., Dumont, J. E., and Vassart, G. (1989) *Science* **244**, 569–572
28. Wilden, U., Hall, S. W., and Kuhn, H. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 1174–1178
29. Lohse, M. J., Benovic, J. L., Codina, J., Caron, M. G., and Lefkowitz, R. J. (1990) *Science* **248**, 1547–1550
30. Minowa, M. T., Minowa, T., Monsma, F. J., Sibley, D. R., and Mouradian, M. M. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 3045–3049
31. Roesler, W. J., Vandenbark, G. R., and Hanson, R. W. (1988) *J. Biol. Chem.* **263**, 9063–9066
32. Lomansey, J. W., Cotecchia, S., Lorenz, W., Leung, W. Y., Schwinn, D. A., Yang-Feng, T. L., Brownstein, M., Lefkowitz, R. J., and Caron, M. G. (1991) *J. Biol. Chem.* **266**, 6365–6369
33. Collins, J. (1981) *Cold Spring Harbor Symp. Quant. Biol.* **45**, 409–416
34. Job, D., Job, C., Mercoyrol, L., and Shire, D. (1991) *Eur. J. Biochem.* **195**, 831–839
35. Proudfoot, N. J. (1989) *Trends Biochem. Sci.* **14**, 105–110
36. Williams, W. L., and Müller, U. R. (1987) *J. Mol. Biol.* **196**, 743–755
37. Stalker, D. M., Shafferman, A., Tolun, A., Kolter, R., Yang, S., Helinsi, D. R. (1981) *ICN-UCLA Symp. Mol. Cell. Biol.* **22**, 113–124
38. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S., and Sharp, P. A. (1986) *Annu. Rev. Biochem.* **55**, 1119–1150
39. Dominski, Z., and Kole, R. (1991) *Mol. Cell. Biol.* **11**, 6075–6083
40. Sehgal, A., Patil, N., and Chao, M. (1988) *Mol. Cell. Biol.* **8**, 3160–3167
41. Smale, S. T., and Baltimore, D. (1989) *Cell* **57**, 103–113
42. Dynan, W. S., and Tjian, R. (1983) *Cell* **35**, 79–87
43. Smale, S. T., Schmidt, M. C., Berk, A. J., and Baltimore, D. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4509–4513
44. Douglas, J., Civelli, O., and Herbert, E. (1984) *Annu. Rev. Biochem.* **53**, 665–715
45. Simon, M. I., Strathmann, M. P., and Gautam, N. (1991) *Science* **252**, 802–808
46. Harrison, J. K., Pearson, W. R., and Lynch, K. R. (1991) *Trends Pharmacol. Sci.* **12**, 62–67